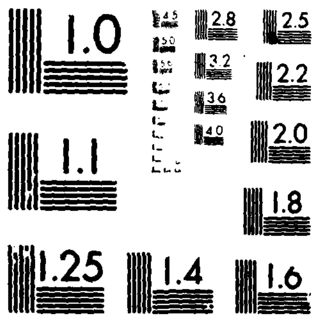
AD-A081 664 FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS F/G 12/1
BOUNDS ON ORDER STATISTICS DIFFERENCES.(U)
OCT 79 S FAHMY, F PROSCHAN AFOSR-78-3678
UNCLASSIFIED FSU-STATISTICS-M525 AFOSR-TR-80-0111 NL

1 OF 1
AD
A081664

END
DATE
FILMED
4-80
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(18) AFOSR TR-80-0111

(4)

LEVEL II

(6) BOUNDS ON ORDER STATISTICS DIFFERENCES,

by

(10) Salwa Fahmy [redacted] Frank Proschan

(14) FSU-Statistics ~~Report~~ M525, TR-78-102-AFOSR
AFOSR ~~Technical~~ Report No. 78-102

(11) October, 1979

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

(16) 2304 (17) A5

(15) Research sponsored by the Air Force Office of Scientific Research, AFSC, USAF, under Grant AFOSR-78-3678.

AMS 1970 Subject Classification: Primary: 62G30.

KEY WORDS: Bounds on order statistic differences, spacings, standard deviation, range, inequalities.

9 Interim rpt.

Bounds of Order Statistics Differences

by

Salwa Fahmy and Frank Proschan

ABSTRACT

A sharp upper bound is obtained for the difference of a pair of order statistics in terms of the sample standard deviation.

1. Results and Derivation. In this note we obtain a sharp upper bound on the difference between a specified pair of sample order statistics in terms of the sample standard deviation. This result falls in the class of results discussed by Wolkowicz and Styan (1979), Arnold (1974), Mallows and Richter (1969), Samuelson (1968), and Thomson (1955).

Let $x_1 \le x_2 \le \ldots \le x_n$ denote the order statistics in a sample of size n. (The underlying distribution plays no role since we confine attention to the one sample only.) Let $\bar{x}$ denote the sample mean $n^{-1}\sum x_i$ and s^2 denote the sample variance $n^{-1}\sum(x_i - \bar{x})^2$.

Given integers k and ℓ $(1 \le k < \ell \le n)$, a sharp upper bound for the difference of order statistics is:

$$x_\ell - x_k \le s/a_{n,k,\ell}, \tag{1}$$

where

$$a^2_{n,k,\ell} = \frac{k(n - \ell + 1)}{n(n - \ell + 1 + k)} . \tag{2}$$

Equality holds in (1) if and only if:

$$x_1 = \ldots = x_k;\ x_{k+1} = \ldots = x_{\ell-1} = \frac{kx_k + (n - \ell + 1)x_\ell}{n - \ell + 1 + k};\ \text{and} \tag{3}$$

$$x_\ell = \ldots = x_n.$$

To prove inequality (1), fix x_k and x_ℓ $(1 \le k < \ell \le n)$ and consider s^2 as a function of the remaining order statistics. We search for the minimum value of s^2 in two stages. First it is clear that s^2 is minimized by requiring that $x_1 = \ldots = x_k$ and $x_\ell = \ldots = x_n$ irrespective of the values assumed by $x_{k+1}, \ldots, x_{\ell-1}$. Now consider s^2 as a function of $x_{k+1}, \ldots, x_{\ell-1}$. It can be shown through differentiation that the minimum value of s^2 is achieved when

$$x_i = \frac{kx_k + (n - \ell + 1)x_\ell}{n - \ell + 1 + k}, \; i = k + 1, \ldots, \ell - 1.$$

The minimum value of s^2 is computed to be

$$a^2_{n,k,\ell}(x_\ell - x_k)^2.$$

Hence (1) follows.

Some special cases of interest are:

Range: $x_n - x_1 \le (2n)^{\frac{1}{2}}s,$ (4)

Symmetric Case: $x_{n+1-k} - x_k \le (\frac{2n}{k})^{\frac{1}{2}}s,$ (5)

Spacing: $x_{k+1} - x_k \le [n/(k^{\frac{1}{2}}(n - k)^{\frac{1}{2}}]s.$ (6)

From (3), the corresponding sample values for which equality is achieved are readily computed.

Inequality (4) was obtained by Thomson (1955).

Remark. We have presented an upper bound on the difference $x_\ell - x_k$ of sample order statistics in terms of the sample standard deviation s (see (1)). Note that no comparable nontrivial lower bound exists for $x_\ell - x_k$, $(1 < k < \ell \le n$ or $1 \le k < \ell < n)$. This is apparent since s can be made arbitrarily large by choice of $x_1(x_n)$ in the first (second) case. However, a sharp lower bound exists for the sample range:

$$x_n - x_1 \ge 2s. \tag{7}$$

See Thomson (1955).

2. Applications. Consider the statistic $(x_\ell - x_k)/s$ which can be used to test for outliers. Inequality (1) shows that this statistic is bounded from above irrespective of the underlying distribution and gives the bound.

In life testing situations the approximate magnitude of s might be (statistically) known from previous experience. From this knowledge, inequality (6) would furnish a conservative estimate of the waiting time between the k^{th} and $k+1^{st}$ failures.

3. Numerical Examples. A few numerical examples may give the reader a clearer picture of the bounds in (1):

Symmetric Case: $n = 15$, $k = 3$: $x_{13} - x_3 \le 3.16s$

$n = 1000$, $k = 3$: $x_{998} - x_3 \le 25.82s$

Spacing: $n = 15$, $k = 9$: $x_{10} - x_9 \le 2.04s$

$n = 1000$, $k = 500$: $x_{501} - x_{500} \le 2.00s$.

Note that the bounds in the latter two cases are very close since the bound depends only on the ratio k/n.

REFERENCES

1) Arnold, Barry C. (1974). Schwarz, Regression, and Extreme Deviance. The American Statistician 28 22-23.

2) Mallows, C. L. and Richter, Donald (1969). Inequalities of Chebyshev Type Involving Conditional Expectations. Ann. Math. Statist. 40 1922-1932.

3) Thomson, George W. (1955). Bounds for the Ratio of Range to Standard Deviation. Biometrika 42 268-269.

4) Wolkowicz, H. and Styan, G. (1979). Extensions of Samuelson's Inequality. The American Statistician 33 143-144.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 80-0111	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) BOUNDS ON ORDER STATISTICS DIFFERENCES		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Salwa Fahmy Frank Proschan		8. CONTRACT OR GRANT NUMBER(s) AFOSR 78-3678
9. PERFORMING ORGANIZATION NAME AND ADDRESS Florida State University Department of Statistics Tallahassee, FL 32306		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, D. C. 20332		12. REPORT DATE October 1979
		13. NUMBER OF PAGES Six
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bounds on order statistic differences, spacings, standard deviation, range, inequalities.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A sharp upper bound is obtained for the difference of a pair of order statistics in terms of the sample standard deviation.

Unclassified

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE